AD_____

Award Number: DAMD17-97-1-7070

TITLE: The Use of a New Technique to Study DNA Methylation in Breast Cancer

PRINCIPAL INVESTIGATOR: Hao-Peng X. Duffy, M.D., Ph.D.

CONTRACTING ORGANIZATION: North Shore University Hospital
Manhasset, New York 11030-3801

REPORT DATE: October 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1999	Annual (1 Oct 98 - 30 Sep 99)

4. TITLE AND SUBTITLE

The Use of a New Technique to Study DNA Methylation in Breast Cancer

5. FUNDING NUMBERS

DAMD17-97-1-7070

6. AUTHOR(S)

Duffy, Hao-Peng X., M.D., Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

North Shore University Hospital
Manhasset, New York 11030-3801

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The mechanisms of breast cancer development are incompletely understood. A reasonable hypothesis is that the evolution of this disease is caused by various genetic, and non-genetic factors. These may occur in multiple steps which eventually precipitate genetic lesions; such as mutations, or deletions. One causative factor may be imbalanced DNA methylation. A new technique, Methylation Differential Display (MDD), has been developed to explore DNA methylation patterns in breast cancer. *BR50* was one of the hypomethylated DNA fragments isolated by MDD. Further study of *BR50* has led to the discovery of a novel gene, *TSP50*. The *TSP50* gene product is homolog to several human serine proteases which indicates that it may encode a serine protease like protein. Northern analysis of sixteen different types of normal human tissues suggest that *TSP50* was highly, and specifically expressed in human testes, which indicates that it might possess a unique biological function(s) in that organ. Methylation status analysis in normal human testes and other tissues showed a correlation between DNA methylation and gene expression. Most importantly, RT-PCR analysis of eighteen paired breast cancer tissues found that, in 28% of the cancer samples, the *TSP50* gene was differentially expressed.

14. SUBJECT TERMS
Breast Cancer Methylation Differential Display technique

15. NUMBER OF PAGES
22

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

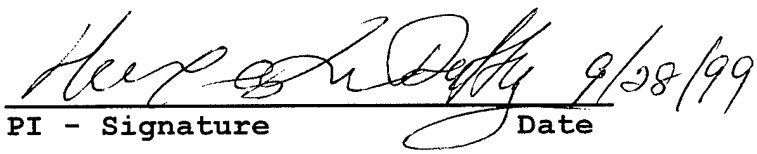
N/A In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

X In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

X In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

X In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.



PI - Signature Date

Table of Contents

The Annual Report for Award Number DAMD17-97-1-7070

The Proposal Title: The Use of a New Technique to Study DNA Methylation in Breast Cancer

The overall goal of the proposal funded by the Department of Defense Breast Cancer Research Program, is to understand the role(s) played by DNA methylation in human breast cancer development. To reach this objective, our specific aims are:

1. to develop and refine a new technique, Methyl Differential Display (MDD), to study the role(s) of DNA methylation in human breast cancer development.
2. to apply this technique to isolate genomic markers which detect altered DNA methylation patterns in breast cancer cells.
3. to search for new types of oncogenes and recessive oncogenes whose expressions are under the control of DNA methylation machinery by differentially methylated genomic DNA fragments (DMGFs).
4. to determine biological function(s) of the candidate genes.
5. to search for their (genomic markers and candidate genes) potential use in clinical diagnosis and prognosis.

In the past year, our studies progress has included the following achievements: 1). To search for a coding region, a DMGF, *BR50*, was used as a probe to screen a human genomic DNA phage library, and a longer DNA fragment, containing at least two exons, was discovered; 2). Based on the sequence information of those two exons, a 700 bp coding region was generated by PCR amplification; 3). Northern analysis demonstrated that the target gene was highly expressed in human testes; 4). The full length gene, *TSP50*, was then acquired from a human testes cDNA library; 5). Homologous analysis revealed that the new gene could be a serine protease; 6). Methylation sensitive southern blots suggest that expression of *TSP50* could be regulated by DNA methylation; and 7). The *TSP50* gene was abnormally activated in some breast cancer patients.

In the following paragraphs, a detailed report of our achievements will be presented. The overall achievements related to the *TSP50* gene have been published in the journal of *Cancer Research* (1, see Appendix).

Materials and Methods

Northern analysis. Two Human Multiple Tissue Northern Blot panels, MTN™, and MTN™ II, were purchased from Clontech Inc. The MTN™ blot contains approximately 2 μg of polyA+ RNA per lane from eight different human tissues (heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas). The MTN™ II blot contains the same amounts of mRNA from an additional eight different human tissues (spleen, thymus, prostate, testes, ovary, small intestine, colon, and peripheral blood leukocyte). The probes labeling and detection were the same as above.

Human complimentary DNA (cDNA) library screening. A human testes λgt11 cDNA library (Human Testis cDNA library, Clontech. Inc.) was used to obtain an intact gene following the instructions of the manufacturer. 2 X 10^5 plaques evenly distributed on 6 plates (150 X 15 mm) were transferred onto Hybond N-membranes. Southern analysis was performed as before. The phage DNA, with human cDNA insert, was purified by the λ Quick! Spin Kit (BIO 101, Inc.) following the instructions of the manufacturer. The individual insert was released by restriction enzyme *Eco* RI cleavage from the phage DNA arms, then subcloned into pUC118 plasmid.

Reverse Transcription-PCR (RT-PCR). Total RNAs were isolated from paired breast cancer and normal tissues by RNA isolation kit, RNA STAT-60 (TEL-TEST, Inc.). The first strand cDNA was synthesized by SuperScript Preamplification system kit (GIBCOBRL, Life Technologies). Oligomers E (ACCAGAGCGTCCAGTGTGTCC, sense) and F (TGGGACTTGATGATCTGAACC, antisense) were used to synthesize the *TSP50* gene. The predicted size was 699 bp. *ß-actin* was used as an internal control whose sense and antisense primers were: 5'-GACGACATGGAGAAGATCTGG-3' and 5'-TGTAGAGGTAGTCAGTCAGG-3'. The predicted size for *ß-actin* was 335 bp. The PCR reaction mixture consisted of cDNA derived from 125 ng of RNA, 10 pmole of sense and antisense primers from both *TSP50* and *ß-actin*, 200 μM of four deoxynucleotide triphosphate, and 0.125 unit of Taq DNA polymerase with reaction buffer (Perkin Elmer) in a final volume of 25 μl. Thirty eight cycles of PCR were carried out. Each cycle of PCR included 30 seconds of denaturation at 95 °C, 60 seconds of annealing at 60 °C, and 60 seconds of extension at 72 °C. The PCR products were separated on a 2% agarose gel.

Results

Search for the coding regions by probe *BR50*. Based on our preliminary analysis, *BR50* was an interesting probe to be used for an adjacent gene search because it was differentially methylated in some breast and ovarian cancer samples tested, even though the sample size was small. *BR50* is a genomic fragment, and consequently the chances that it encodes a polypeptide are slim because a majority of human genomic DNA sequences are non-coding sequences. As a result, we decided that the first step in searching for a gene should be to isolate a longer DNA piece from a human genomic DNA library, with the hope that it may contain exon(s). It is well known that DNA methylation sites can be near genes, therefore, we decided to screen a human placenta genomic phage library, EMBL3 SP6/T7 (Clontech, Inc.), where the insertion sizes were relatively small (9 to 28 kbp). For this purpose, 2 X 10^6 plaques derived from the library were screened by probe *BR50*. As a result, a 17 kbp length clone was isolated. To obtain the sequence information, the DNA clone was released from the phage DNA arms by restriction enzyme *Sst* I cleavage which generated eight DNA fragments. All eight fragments were subcloned into pUC118 plasmids. The fragments smaller than 1 kbp were completely sequenced, while the fragments larger than 1 kbp were partially sequenced from both the plasmid and insert junctions. A homolog search of the NIH GeneBank revealed two exons, *BR50-44* and *BR50-45*, which contained 112 and 132 nucleotides, respectively. Both exons encoded polypeptides which were about 50% identifiable to several mammalian proteases, such as serine proteases and tryptases. The

BR50-45 sequence was found 142 bps up stream of the *BR50* sequence.

Since we did not know the exact positions of the two exons in the 17 kbp fragment, it was possible that other exon(s) might lay between them. To gain a longer coding sequence, we designed four oligomers (A, B, C, D) based on both exon's sequence information to perform PCR. Oligomer A (5'-CCTGGATGGTCAGCGTG-3') and B (5'-CTGGGAGGCAATGATGGT-3'), which were on the complimentary strand, were based on the sequence information of *BR-44*; and C (5'-CTGGAGAGCCCTTGGTCT-3') and D (5'-CAGTGTTGGTAGGAGGAG-3'), which were on the complimentary strand, were based on the sequence information of *BR-45*. A strategy using four different combinations of oligomer pairs was employed to perform PCR by utilizing the Human Universal cDNA Library Panel (Clontech Inc.). A PCR product which was about 700 bps in length was generated from one oligomer combination (A/D, 5'-CCTGGATGGTCAGCGTG-3'/5'-CAGTGTTGGTAGGAGGAG-3'). This PCR product was directly sequenced. Combining the sequence information of the PCR product and the two exons, we obtained a cDNA fragment which contained 974 bps. The DNA homolog search of the NIH GeneBank revealed again, that it coded for a protease like protein, and the overall identity was approximately 40%.

The candidate gene is highly, and specifically expressed in human testes. To obtain a full length cDNA, it is critical to use the right cDNA library where the gene of interest is expressed. Thus, two Human Multiple Tissue Northern Blot panels, MTN™ and MTN™ II, containing 16 different tissue mRNAs (Clontech Inc.), were used to test the expression of the candidate gene by using the 700 bp cDNA PCR product as a probe. The results showed that there were no visible transcripts of this gene in the eight mRNAs (heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas) included in the MTN™ panel (data not shown). In the MTN™ II panel, a 1.7 kb band was heavily hybridized by the probe exclusively in the testes mRNA as compared to the control probe, which was the human *rab6* gene (Fig. 1). These results suggested that the gene that we were searching for is a tissue specific gene. We have named the gene *TSP50* (Testes Specific Protease). At this moment, the gene's biological function(s) in human testes remains unknown.

Isolation of the full length *TSP50* gene from a human testes cDNA library. Since *TSP50* is highly expressed in human testes, to search for the full length *TSP50* gene, a human testes cDNA library (Human Testis cDNA library, Clontech. Inc.) was screened by the 700 bp *TSP50* cDNA sequence. A cDNA clone containing the probe sequence was isolated. Sequence analysis suggests that this fragment encodes a protein with 385 animo acids (Fig. 2.). There is a stop codon located at the 117th bp up stream of the initial translation site, and there is a 125 bp untranslated region before a polyadenosine sequence. These results imply that a full length gene has been obtained. It is also notable that the *BR50* probe sequence is located at the 3' end of the gene, and is only 17 bps down stream from the polyadenosine adding signaling site. The exons, *BR50-44* and *BR50-45*, encode animo acids from 103 to 157 and 308 to 385, respectively. The 3'- untranslated region before the polyadenosine site is also included in the sequence of *BR50-45*.

DNA methylation status of the *TSP50* gene in human testes and other normal tissues. Our studies have proven that *TSP50* is a tissue specific gene, and the methylation patterns in its

3'- region were altered in some breast and ovarian cancers. It is also known that many tissue specific genes are methylated, and this methylation may regulate their expression (2-4). To explore the possible relationship between *TSP50* gene expression and DNA methylation in different normal human tissues, southern analysis was performed. The normal tissues tested included the testes, where *TSP50* was expressed, and bladder, blood, breast, colon, lung, kidney, placenta, and ovary samples, where *TSP50* was apparently not expressed. To perform the southern analysis, *BR50* was used as a probe. DNAs isolated from the nine tissues were digested by *Msp* I, and *Hpa* II, which is an isoschizomer of the *Msp* I enzyme and the most popular enzyme used to study DNA methylation patterns (5). *Hpa* II digestion showed that in the testes DNA, two bands, probably released from each allele by enzyme cleavage, were hybridized by the probe. However, in the other tissues' DNAs, the corresponding bands were either not hybridized, or hybridized to a much smaller degree (Fig. 3a). For *Msp I* cleavage, both bands were released in different tissues to various extents (Fig. 3b). Both blots utilized a genomic fragment which did not detect differential DNA methylation as a control to determine complete enzymatic digestion (Fig. 3). These results demonstrated that the *TSP50* gene was differentially methylated in various human tissues. In general, DNA demethylation in the testes is correlated with high levels of gene expression. Conversely, DNA methylation is correlated with gene silencing in the bladder, blood, breast, colon, lung, kidney, placenta, and ovary tissues.

Comparison of TSP50 product sequence with other serine proteases. The sequence analysis revealed that the *TSP50* gene encodes a protein which shares approximately 40% identity with mammalian serine proteases. Figure 2 compares the *TSP50* animo acid sequence with 7 other serine proteases including prostasin (6), plasma kallikrein (7), coagulation factor XI (8), ß-tryptase (9), hepsin (10), plasminogen (11), and acrosin (12). Proteolytic enzymes dependent on a serine residue for catalytic activity are widespread and very numerous. Serine proteases are found in viruses, bacteria, and eukaryotes, and they include exopeptidases, endopeptidases, oligopeptidases, and omega peptidases. Over 20 families of serine peptidases are recognized (13,14), and grouped into clans that may have common ancestors. The peptidases of chymotrypsin, subtilisin, and carboxypeptidase C clans have in common a "catalytic triad" of three amino acids: Serine (Ser, nucleophile). Aspartate (Asp, electrophile), and Histidine (His, base). However, there are some serine peptidases that have distinctive mechanisms of action without the classic Ser, His, Asp triad. The multiple sequence alignment for tsp50p showed that it contains triad His^{153} and Asp^{206}. However, the Ser at position 310 has been replaced by threonine (Thr) (Fig. 2). The corresponding nucleotides for coding Thr in the *TSP50* gene were ACT, while one base pair switch, such as C to G, will result in a Ser codon, AGT. As a result, one may wonder whether this change was caused by a point mutation happening in the cells, or an error in DNA sequencing. Based on our experience, these assumption are unlikely since the DNA fragment containing this codon was isolated from two individual libraries, the human placenta genomic phage library and testes cDNA library, and the ACT codon was verified by DNA sequencing in both fragments. This implies that the Ser was replaced by Thr in the predicted Ser triad of tsp50p. However, Thr and Ser residues are structurally similar (Thr has an extra methyl group as compared to Ser). Both Thr and Ser contain the HO group that is critical for enzymatic catalysis. In addition, the Thr residue in tsp50p was surrounded with conserved residues including a crucial glycine (Gly) (Fig. 2). Usually the linear order of catalytic-site residues,

clusters of conserved amino acids around the catalytic residues, are important factors to classify a protease (13,14). Therefore, tsp50p could be a new type of serine protease, possibly with a distinctive mechanism of action.

***TSP50* was differentially expressed in some breast cancer tissues.** Preliminary results demonstrated that *TSP50* was differentially methylated in 40% of the breast cancer tissues tested. This suggested that it could also be differentially expressed in cancer. To test this possibility, RT-PCR was carried out to determine *TSP50* expression levels in eighteen paired breast cancer tissues. Our findings showed that *TSP50* PCR products were generated in five tumor tissues, while in their normal controls, they were not visible relative to the control gene, *ß-actin* (Fig. 4). Products generated from the five patients were gel purified and sequenced. DNA sequence analysis confirmed that the PCR products synthesized were the *TSP50* gene (data not shown). Therefore, among the eighteen paired samples tested, 28% of the tissues expressed the *TSP50* gene. At this moment, we can not answer the question of whether activation of the *TSP50* gene in cancer is a consequence, or a causal factor of neoplastic growth. To find the truth, it will be necessary to perform in vitro cellular transformation and in vivo tumor induction assays.

Discussions

Since DMGF generated by the MDD technique could be close to gene of interest our first step in processing the gene search was to screen a human genomic phage library. A 17 kbp DNA fragment was isolated and sequence analysis suggested that this fragment contained at least two exons which were homolog to mammalian proteases. The exons' sequencing information led to the discovery of a 974 bp gene fragment from a human cDNA library panel by PCR amplification. To obtain a full length gene, a northern evaluation on sixteen different types of human RNAs was performed. The results demonstrated that the target gene was specifically expressed in human testes tissue. This information secured the isolation of an intact gene, *TSP50*, by screening a human testes cDNA library. The sequence analysis revealed that the *TSP50* gene most likely encodes a serine protease.

It is well known that proteases of all major classes (i.e., serine, aspartic, cysteine, threonine, and metalloproteinases) are linked with various malignancies, especially those exhibiting the metastasis phenotype. For examples, prostate specific antigen (PSA) is a kallikrein-like serine protease that is utilized as a clinical marker for the diagnosis and staging of prostate cancer where its preferential expression in prostate epithelial cells is increased (15,16). In addition, the matrix metalloproteinases (MMPs) have been repeatedly implicated in metastasis (17-19). Since the *TSP50* gene product (tsp50p) could be a serine protease, and was activated in breast cancer cells, it is logical to proceed to the next step and asks whether this gene plays a role in breast cancer invasiveness?

It is common knowledge that many tissue specific genes' expression is regulated by DNA methylation which usually modify the promoter, or sometimes, 3'- regions (2-4, 20). Our preliminary results, although only obtained from analyzing the DNA methylation status of the 3' flanking region of the gene, have proven that *TSP50* is one of those tissue specific genes. It will be interesting to discover whether the gene's promoter region is also

methylated when the corresponding sequence information is available. The *Hpa* II and *Msp* I methylation sensitive southern analysis of the *TSP50* gene's 3'- region demonstrated that, in *Hpa* II digested DNAs, probe *BR50* hybridized two bands in the testes tissue, but none in the other samples. The lower band, which was the same size as the probe, represented the unmethylated DNA pattern, while the upper band obviously contained the internal *Hpa* II recognition site(s) which remained methylated. In *Msp* I digested DNAs, the upper band was dominant in most tissues, while in the testes, the lower band was dominant. These results suggest that the GGCCGG end of *BR50* was methylated in other tissues, but not in the testes. The DNA methylation patterns observed in both blots are probably allelic oriented. It seems that DNA hypomethylation was accompanied by the gene's expression in the testes, and conversely, DNA hypermethylation was accompanied by the gene's silencing in other tissues. The correlation between DNA methylation and gene expression provided additional proof that DNA methylation could be an important mechanism in governing the genes' expression in various differentiated human cells.

In addition, the differential expression of the *TSP50* gene has been tested in eighteen paired breast cancer biopsies. Our findings have shown that this gene was activated in five cancer samples. This finding indicates that this novel gene's expression is related to breast cancer progression. In the near future, more samples from different types of cancer will be examined, and the possibility that the *TSP50* gene product might be one of the factors that stimulate human cancer will be further explored.

GenBank accession numbers. *BR50*: U78781, *TSP50*: AF100707.

Figures and Figure Legends.

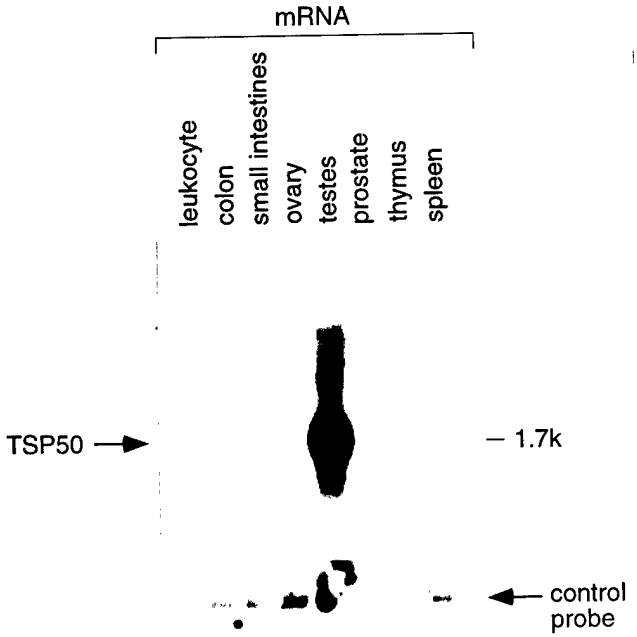


Fig. 1. Northern blot results analyzed by a fragment of the *TSP50* gene. The Human Multiple Tissue Northern Blot panel, MTN™ II, containing 8 different tissues' mRNA (Clontech Inc.) was tested to determine the expression levels of the gene. Compared to the control, the human Rab6 gene, which was evenly expressed in all tissues, *TSP50* was highly expressed in the testes tissue, but not in the others. No single tissue (total eight) in another Human Multiple Tissue Northern Blot panel was hybridized with the probe (data not shown).

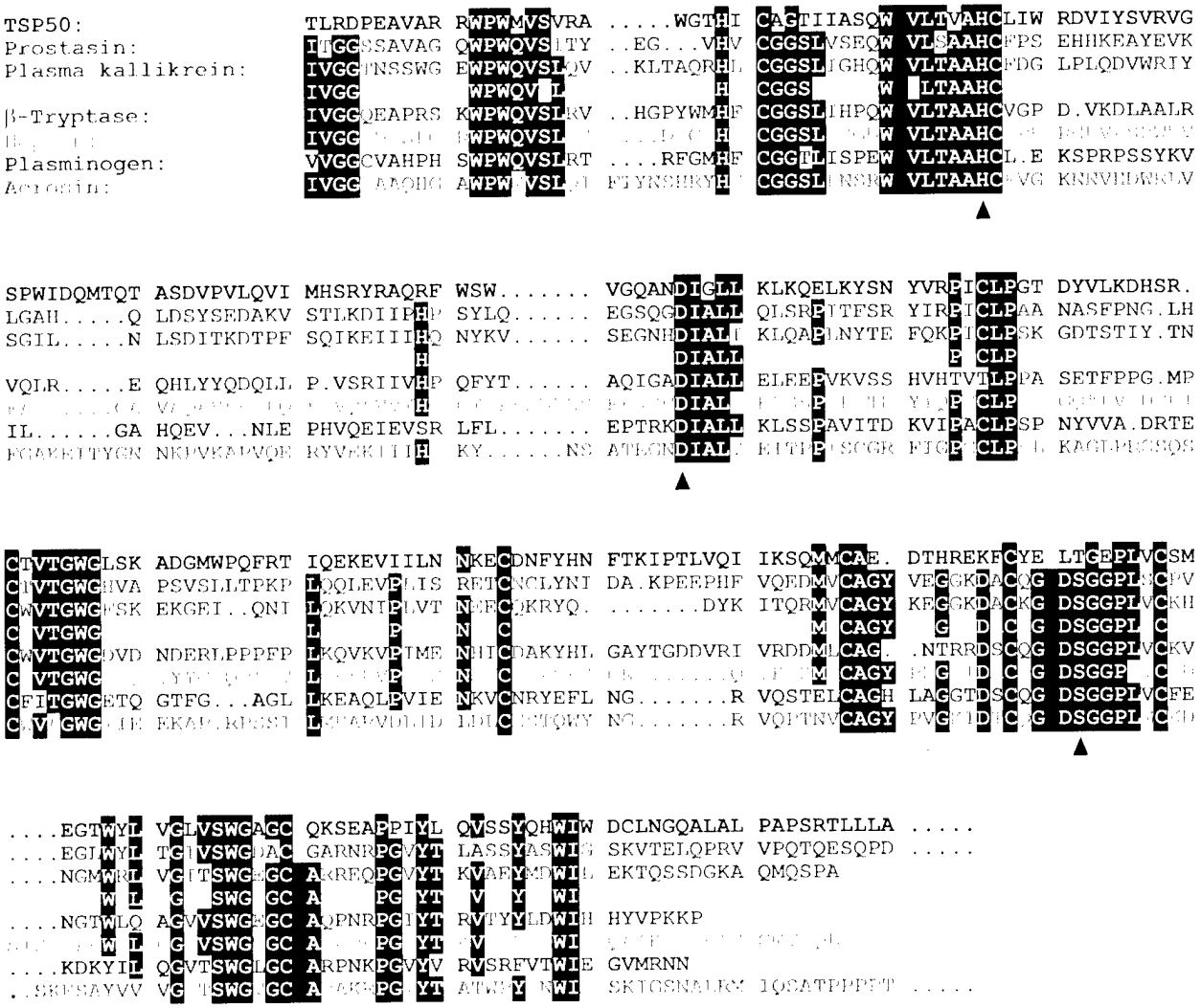


Fig. 2. Comparison of the deduced *TSP50* animo acid sequence with other serine proteases. The amino acid sequences of these serine proteases correspond to the mature form of ß-trypsin or the catalytic chains of acrosin, prostasin, plasma kallikrein, coagulation factor XI, serine protease hepsin, and plasminogen. Amino acid residues that are highly conserved are shaded, and the catalytic triad of histidine, aspartic acid, and serine are indicated by triangles. Dots represent gaps to bring the sequences to better alignment.

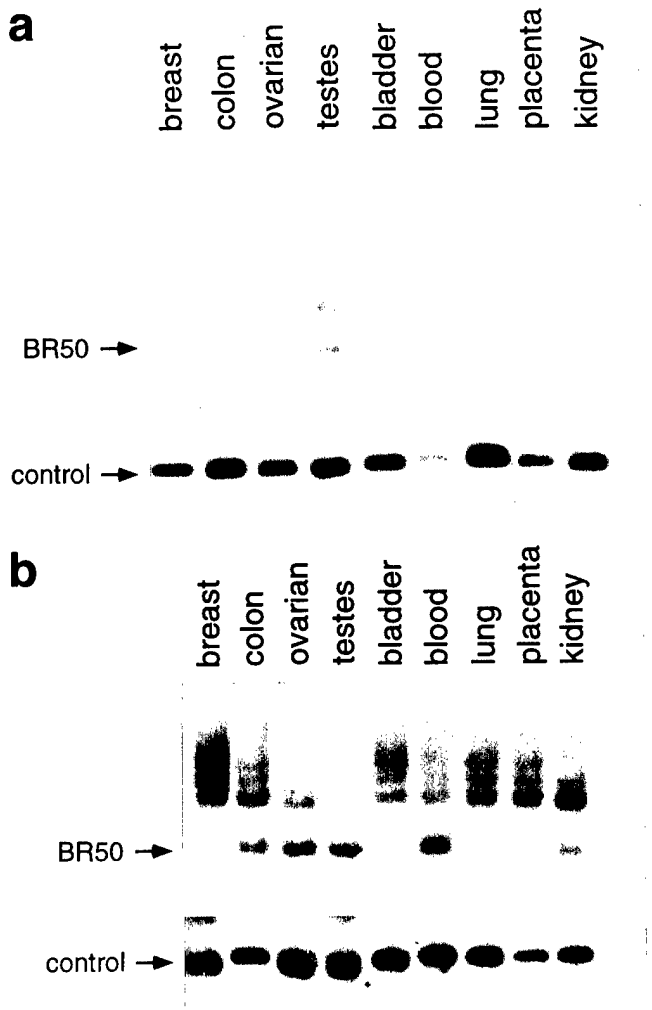


Fig. 3. DNA methylation status of the *TSP50* gene in nine normal human tissues examined by southern blot. In a and b, the results obtained from *Hpa* II and *Msp* I digestion, respectively. 6 µg of DNA isolated from each tissue were cleaved by the enzymes and subjected to southern analysis by probe *BR50*. a. The results show that bands which are approximately 1 kbp and 2 kbp in length were released by *Hpa* II only in the testes tissue. b. A 2 kbp band was released by *Msp* I in most tissues. In a and b. The control probe hybridizes a single band in each tissue's DNA, which provides proof of complete enzymatic digestion.

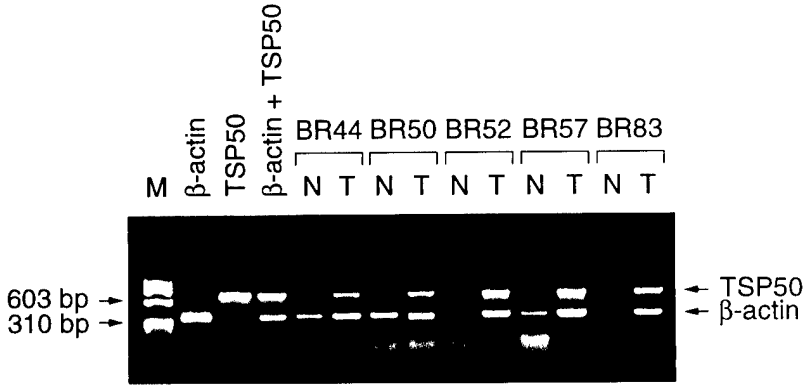


Fig. 4. The results of RT-PCR for differential expression of *TSP50* in five out of eighteen breast cancer and normal control tissues tested. Lane MW represents *Hae* III ø174 markers in base pairs. The *ß-actin* and *TSP50* lanes served as positive controls, and were generated by using cDNA prepared from testes tissue RNA. Lane *ß-actin+TSP50* contains simultaneously generated *TSP50* and *ß-actin* from testes cDNA. The number for each patient tested is listed above the bracket. T and N represent tumor and normal tissues. The result shows that the *TSP50* gene was abnormally activated in approximately 30% of breast cancer patients.

References:

1. Yuan, L. M., Shan, J. D., De Risi, D., Broome, J., Lovecchio, J., Gal, D., Vinciguerra, V., and Xu, H. P., 1999. Isolation of a novel gene, *TSP50*, by a hypomethylated DNA fragment in human breast cancer. Cancer Res., 59, 3215-3221.
2. Bird, A. Gene number, noise reduction and biological complexity. Trends Genet., *11*: 94-100, 1995.
3. Kass, S. U., Pruss, D., and Wolffe, A. P. How does DNA methylation repress transcription? Trends Genet., *13*: 444-449, 1997.
4. Siegfried, Z., and Cedar, H. DNA methylation: a molecular lock. Curr. Biol., *7*: R305-R307, 1997.
5. Nelson, M., and McClelland, M. Site-specific methylation: effect on DNA methylation methyltransferases and restriction endonucleases. Nucl. Acids Res., *19*: 2045-2071, 1991.
6. Yu, J. X., Chao, L., and Chao, J., 1995. Molecular cloning, tissue-specific expression, and cellular location of human prostasin mRNA. J. Biol. Chem. 270, 13483-13489.
7. Chung, D. W., Fujikawa, K., McMullen, B. A., and Davie, E. W., 1986. Human plasma prekallikrein, a zymogen to a serine protease that contains four tandem repeats. Biochemistry 25, 2410-2417.
8. Fujikawa, K., Chung, D. W., Hendrickson, L. E., and Davie, E. W., 1986.Amino acid sequence of human factor XI, a blood coagulation factor with four tandem repeats that are highly homologous with plasma prekallikrein. Biochemistry 255, 2417-2424.
9. Vanderslice, P., Ballinger, S. M., Tam, E. K., Goldstein, S. M., Craik, C. S., and Caughey, G. H., 1990.Human mast cell tryptase: multiple cDNAs and genes reveal a multigene serine protease family. Proc. Natl. Acad. Sci. U. S. A. 87, 3811-3815.
10. Leytus,, S. P., Loeb, K. R., Hagen, F. S., Kurachi, K., and Davie, E. W., 1988.A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain expressed by human liver and hepatoma cells. Biochemistry 27, 1067-1074.
11. Forsgren, M., Raden, B., Israelsson, M., Larsson, K., and Heden, L.-O., 1987. Molecular cloning and characterization of a full-length cDNA clone for human plasminogen. FEBS Lett. 213, 254-260.
12. Adham, I. M., Klemm, U., Maier, W.-M., and Engel, W., 1990. Molecular cloning of human preproacrosin cDNA. Hum. Genet. 84. 125-128.
13. Barrett, A. J., 1994. Classification of peptidases. Methods in Enzymol. 244, 1-14.
14. Rawlings, N. D., and Barrett, A. J., 1994. Families of serine peptidases. Meth. Enzymol. 244, 19-31.
15. Wang, M. C., Valenzuela, L.A., Murphy, G. P., and Chu, T. M., 1979. Purification of a human prostate specific antigen. Invest. Urol. 17, 159-163.
16. Yoshihito, B., Wang, M. C., Watt, K. W. K., Loor, R., and Chu, T. M., 1984. The proteolytic activity of human prostate-specific antigen. Bioch. & Biophy. Res. Commun. 123, 482-488.
17. Chambers, A. F., and Matrisian, L. M., 1997. Changing views of the role of matrix metalloproteinases in metastasis. J. Natl Cancer Inst. 89, 1260-1270.
18. Powell W. C., and Martrisian, L. M., 1996. Complex roles of matrix metalloproteinases in tumor progression. In: Gunthert U. Birchmeier W. Editors. Attempts to understand metastasis formation I: metastasis-related molecules. Berlin: Spring-

Verlag, p1-21.

19. Birkedal-Hansen, H., Moore, W. G., Bodden, M. K., Windsor, L. J., Birkedal-Hansen, B., DeCarlo, A., et al., 1993. Matrix metalloproteinases: a review. Crit. Rev. Oral Biol. Med. 4, 197-250.

20. Graff, J. R., Herman, J. G., Myohanen, S., Baylin, S. B., and Vertino, P. M. Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in *de novo* methylation. J. Biol. Chem., 272, 22322-22329, 1997.

Appendix

[CANCER RESEARCH 59, 3215–3221, July 1, 1999]

Isolation of a Novel Gene, *TSP50,* by a Hypomethylated DNA Fragment in Human Breast Cancer[1]

Liming Yuan, Jidong Shan, Dwight De Risi, John Broome, John Lovecchio, David Gal, Vincent Vinciguerra, and Hao-peng Xu[2]

Molecular Oncology, Hematology/Oncology Medicine [L. Y., J. S., V. V., H-p. X.], Department of Surgery [D. D. R.], Department of Laboratories [J. B.], and Division of Gynecologic Oncology [J. L., D. G.], North Shore-Long Island Jewish Health System, New York University School of Medicine, Manhasset, New York 11030

ABSTRACT

A novel gene, testes-specific protease 50 (*TSP50*), was isolated from a human testes cDNA library by using a genomic DNA probe, *BR50*. *BR50* was isolated by a modified representational difference analysis (RDA) technique due to its hypomethylated feature in a breast cancer biopsy. This altered DNA methylation status was also detected by *BR50* in other breast and some ovarian cancer tissues. The *TSP50* gene product is a homologue to several human proteases, which indicates that it may encode a protease-like protein. Northern analysis of 16 different types of normal human tissues suggests that *TSP50* was highly and specifically expressed in human testes, which indicates that it might possess a unique biological function(s) in that organ. Methylation status analysis in normal human testes and other tissues showed a correlation between DNA methylation and gene expression. Most importantly, reverse transcription-PCR analysis of 18 paired breast cancer tissues found that in 28% of the cancer samples, the *TSP50* gene was differentially expressed. The possibility that *TSP50* may be an oncogene is presently under investigation.

INTRODUCTION

Abnormal DNA methylations (hypomethylation or hypermethylation) have been linked to various human diseases including cancers (1–8). Because methylated DNA sites are usually close to genes (9–13), searching for differentially methylated DNA fragments in cancer could pinpoint genes of interest. Consequently, a modified RDA technique using human breast cancer biopsies as starting material was used to search for differentially methylated DNA fragments. Unlike traditional RDA[3] (14), to perform modified RDA, the restriction enzyme *Msp*I, which is sensitive to the methylated GC-rich sequence GGC^{m}CGG (15–18), was used as a master enzyme to cleave genomic DNAs for amplicon preparation. *Msp*I is a relatively frequent cutting enzyme (19), which, when used alone, produces an amplicon of high complexity that can cause unsuccessful subtractive hybridization. Hence, a second restriction enzyme, or "partner enzyme," has been incorporated into the technique. The amplicons can only be made by PCR from DNA fragments with both ends cut by the *Msp*I enzyme.

As a result of using this modified technique, two DNA fragments, *BR50* and *BR254,* were isolated that detected DNA hypomethylation in breast cancer. Additional studies verified that both fragments also detected hypomethylation in ovarian cancer, and *BR254* was amplified in 1 of 10 breast cancer biopsies. On the basis of these findings, we considered both fragments good candidates to search for genes that might be related to various malignancies. This report will focus on presenting the detailed studies related to *BR50*, which covers its own isolation as a differentially methylated DNA fragment, to its utilization in the discovery of a novel gene, *TSP50.*

Investigation of the *TSP50* gene has found that it encodes a protease-like protein. Northern analysis of multiple human tissue RNA expression panels showed that *TSP50* is a tissue-specific gene, which was heavily expressed in human testes. There were almost no visible amounts of *TSP50* transcript displayed in the other 15 types of human tissues in the panels. This result indicates that the *TSP50* gene holds a special physiological function(s) in human testes. The DNA methylation status of the downstream region of the gene in normal human testes and eight other tissues was also examined. Apparently, DNA methylation silences the *TSP50* gene expression in those eight normal tissues, whereas DNA demethylation in human testes could be a key element responsible for gene expression. Furthermore, RT-PCR was performed to examine differential expression in breast cancer and matched normal control tissues. We found that ~28% of the cancer samples tested expressed the *TSP50* gene, whereas the corresponding controls did not. Whether there is a relationship between gene expression and cancer development is presently under investigation.

MATERIALS AND METHODS

DNA from Human Cancer Biopsies. Dissected human breast and ovarian cancer tissues (tumor and matched normal) were immediately frozen in liquid nitrogen and stored at −70°C. DNAs were isolated from those tissues by the phenol extraction method (20).

Modified RDA. One to two μg of DNA (tester) isolated from human breast cancer biopsies and their matched normal DNA (driver) were cleaved with *Msp*I (20 units/μl; Boehringer Mannheim) and *Mse*I (20 units/μl; New England Biolabs, Inc.) in a 50-μl reaction for 3 h. To prepare tester and driver master amplicons, the *Msp*I- and *Mse*I-digested tester and driver genomic DNAs were ligated to 1.5 μg of MSA24-mer and 0.75 μg of MSA12-mer (Table 1); these were the first pair of oligonucleotide linkers that only recognize the ends generated by *Msp*I. The procedures for amplicon preparation were performed as described (14). The DNA amplicons were then purified by phenol, phenol/chloroform extraction. To remove the first set of linkers from the driver amplicon, 80 μg of driver amplicon DNA were digested with the *Msp*I enzyme (10 units/μl). To change the tester master amplicon DNA linkers, 5 μg of tester master amplicon were digested with *Msp*I (20 units/μl) and ligated to 0.6 μg of MSB24-mer and 0.3 μg of MSB12-mer (Table 1); these were the second set of oligonucleotide linkers. Subtractive hybridization was performed as described (14). The first round of difference products (DP1) were amplified as described (14). To prepare the second round of subtractive hybridization, 3 μg of DP1 were digested with the restriction endonuclease *Msp*I (20 units/μl). To put a new set of linkers on DP1, 0.1 μg of DP1 was mixed with 0.6 μg of MSC24-mer and 0.3 μg of MSC12-mer (Table 1). Another round of subtractive hybridization/PCR amplification was repeated. The second round of difference products (DP2) usually contained several individual DNA fragments when electrophoresed on a 2% agarose gel. The individual DNA fragments were purified by DNA gel extraction kit (Qiagen, Inc.) and subcloned into pUC118 vector, which was linearized by the restriction endonuclease *Acc*I and transformed into *Escherichia coli* (*DH5α*). Twelve cloned inserts were chosen to be amplified, from which different-sized probes were selected for master amplicon Southern blot. The candidate probes were then further tested by human genomic DNA Southern blot.

Amplicon DNA Southern Blot. The first round of positive probe screening was performed with amplicon DNA Southern blots. Non-Radiation South-

Received 10/28/98; accepted 5/3/99.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported by Grant DAMD17-97-7070 from the Department of Defense Breast Cancer Research Program, Cancer Research Grant 09289 from North Shore-Long Island Jewish Health System, and the Greenberg Foundation.

[2] To whom requests for reprints should be addressed, at Molecular Oncology, Research Building, North Shore-Long Island Jewish Health System, 350 Community Drive, Manhasset, NY 11030.

[3] The abbreviations used are: RDA, representational difference analysis; RT-PCR, reverse transcription-PCR; DP, difference product; TSP, testes-specific protease.

Table 1 *Sequence of the oligonucleotides for PCR amplification*

Primer	Sequence
MSA24	5′-CTCGTCGTCAGGTCAGTGCTTCAC-3′
MSA12	5′-CGGTGAAGCACT-3′
MSB24	5′-TAGAGCCACGTAGCTGCTGTAGTC-3′
MSB12	5′-CGGACTACAGCA-3′
MSC24	5′-ACCGTGGACTGGATAGGTTCAGAC-3′
MSC12	5′-CGGTCTGAACCT-3′

ern Blot and Detection kits (Genius) were purchased from Boehringer Mannheim. Probe labeling and detection followed the instructions of the manufacturer. Two to three μg of tester and driver amplicon DNA were electrophoresed on a 2% agarose gel and blotted to positively charged nylon membranes (Boehringer Mannheim). For prehybridization, the membranes were placed at 68°C for 2–4 h in solutions containing 6× SSC, 5× Denhardt's solution, 0.5% SDS, 0.1 M EDTA, and 50 μg/ml of salmon sperm DNA. Under the same conditions, the probes were added and hybridized to the membranes overnight. The membranes were then rinsed three times with 2× SSC, 1× Blot wash (12 mM Na_2HPO_4, 8 mM NaH_2PO_4, 1.4 mM $Na_4P_2O_7$, and 0.5% SDS) at 68°C and further washed three times (30 min each) with the same buffer at 68°C. Next, the membranes were equilibrated in buffer A (100 mM Tris·HCl, 150 mM, pH 7.5) and transferred into buffer B (2% block reagent in buffer A), which was incubated at room temperature for 1 h. The membranes were then washed 2 times for 15 min with buffer A and equilibrated in buffer C (100 mM Tris·HCl, 100 mM NaCl, and 10 mM $MgCl_2$). Before the membranes were exposed on Kodak X-OMAT film for 1 h, they were rinsed in lumi-P530 for 1 min and kept in a plastic sheet protector.

Genomic DNA Southern Blot. Genomic DNAs were digested with a desired restriction enzyme (20 units/μl) and electrophoresed on 1.5% agarose gels, which were then transferred to Hybond N^- membranes (Amersham). These membranes were exposed to UV light to immobilize the DNA. Probes for the Southern blot were labeled with High Prime DNA labeling kits (Boehringer Mannheim) following the instructions of the manufacturer. The procedure for hybridization and blot wash were the same as in the Amplicon DNA Southern blot section.

Northern Analysis. Two Human Multiple Tissue Northern blot panels, MTN and MTN II, were purchased from Clontech, Inc. The MTN blot contains ~2 μg of poly(A)$^+$ RNA per lane from eight different human tissues (heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas). The MTN II blot contains the same amounts of mRNA from an additional eight different human tissues (spleen, thymus, prostate, testes, ovary, small intestine, colon, and peripheral blood leukocyte). The labeling and detection of the probes were the same as above.

DNA Sequence and Chromosome Assignment. The pUC118 plasmid containing the candidate DNA fragment was sequenced using the Ampli-Cycle sequencing kit (Perkin-Elmer), under conditions described by the manufacturer. Chromosome assignment for the candidate DNA fragment was determined by genomic Southern blot of the *Hin*dIII digested monochromosomal human/rodent somatic cell hybrid mapping panel #2 (NIGMS Human Genetic Mutant Cell Repository) while it was used as a probe. Fine chromosome mapping was performed with GeneBridge 4 Radiation Hybrid Panel (Research Genetics, Inc.) by PCR amplification (21).

Human Genomic DNA Library Screening. A human placenta genomic phage library, EMBL3 SP6/T7 (Clontech, Inc.), was used for cloning a longer genomic fragment containing the candidate fragment. Phage infection procedure was based on the instructions supplied by the manufacturer. Plaques (2×10^6) were evenly distributed on 20 plates (150 × 15 mm), then transferred onto Hybond N^- membranes. The treatment of the membranes, preparation of the probe, and the blot wash were the same as that described in the Genomic DNA Southern blot section. The phage DNA, with human DNA insert, was purified by the Lambda TRAP Plus kit (Clontech, Inc.) following the instructions of the manufacturer. The individual insert was released by restriction enzyme *Sst*I cleavage from the phage DNA arms, then subcloned into pUC118 plasmid.

Human cDNA Library Screening. A human testes λgt11 cDNA library (Human Testis cDNA Library; Clontech, Inc.) was used to obtain an intact gene following the instructions of the manufacturer. Plaques (2×10^5) evenly distributed on six plates (150 × 15 mm) were transferred onto Hybond N^- membranes. Southern analysis was performed as before. The phage DNA, with human cDNA insert, was purified by the λ Quick! Spin kit (BIO 101, Inc.) following the instructions of the manufacturer. The individual insert was released by restriction enzyme *Eco*RI cleavage from the phage DNA arms, then subcloned into pUC118 plasmid.

RT-PCR. Total RNAs were isolated from paired breast cancer and normal tissues by RNA isolation kit, RNA STAT-60 (TEL-TEST, Inc.). The first-strand cDNA was synthesized by SuperScript Preamplification system kit (Life Technologies, Inc.). Oligomers E (ACCAGAGCGTCCAGTGTGTCC, sense) and F (TGGGACTTGATGATCTGAACC, antisense) were used to synthesize the *TSP50* gene. The predicted size was 699 bp. *β-actin* was used as an internal control, the sense and antisense primers of which were 5′-GACGACATG-GAGAAGATCTGG-3′ and 5′-TGTAGAGGTAGTCAGTCAGG-3′. The predicted size for *β-actin* was 335 bp. The PCR reaction mixture was comprised of cDNA derived from 125 ng of RNA, 10 pmol of sense and antisense primers from both *TSP50* and *β-actin*, 200 μM of four deoxynucleotide triphosphate, and 0.125 unit of *Taq* DNA polymerase with reaction buffer (Perkin-Elmer) in a final volume of 25 μl. Thirty-eight cycles of PCR were carried out. Each cycle of PCR included 30 s of denaturation at 95°C, 60 s of annealing at 60°C, and 60 s of extension at 72°C. The PCR products were separated on a 2% agarose gel.

RESULTS

Isolation of Hypomethylated Sequences from Human Breast Cancer Biopsies. The DNAs isolated from three paired human breast cancer biopsies (tester) and surrounding normal tissues (driver) were cleaved with the *Msp*I and *Mse*I enzymes. The tester and driver amplicons with both ends cleaved by *Msp*I were selectively prepared by PCR amplification (see "Materials and Methods"). After two rounds of DNA hybridization/subtraction and PCR amplification, individual fragments (DP2) were isolated from two breast cancer patients (see "Materials and Methods"). The DP2 fragments were subcloned into the pUC118 vector, and the inserts were amplified by PCR. Twelve different-sized inserts were selected from each modified RDA and used as probes for the master amplicon Southern blot. Two probes, *BR50* and *BR254*, isolated from two patients, were identified as candidate probes for additional study. In this report, we focus on presenting the work that has been done by probe *BR50*.

Probe *BR50* was selected from the DP2 isolated from breast cancer patient no. 14's biopsy by the modified RDA technique (Fig. 1*a*) because it hybridized a band of much greater intensity in the tester

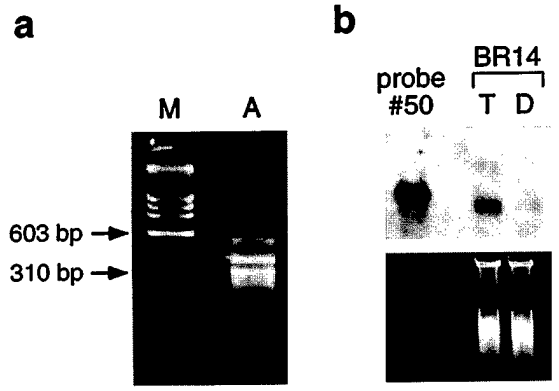


Fig. 1. *a*, the agarose gel electrophoresis of the final difference products isolated by a modified RDA technique from the breast cancer biopsy of patient no. 14. *Lane M, Hae*III ϕ174 DNA size markers in bp. *Lane A*, DNA fragments from which *BR50* was isolated. *b*, the amplicon Southern Blot results for *BR50*. *Left lane*, probe *BR50* hybridizes to itself as a positive control. In *Lanes T* and *D*, probe *BR50* hybridizes with 2 μg of tester (tumor) and driver (normal) master amplicon DNA prepared from breast cancer patient no. 14. The tester master amplicon DNA displays a much heavier hybridized band than the driver master amplicon DNA. The picture below the Southern blot is the tester and driver amplicon agarose gel electrophoresis before transferring it onto the blot membrane, which served as the DNA quantitative control.

amplicon than that in the driver amplicon (Fig. 1*b*). To confirm the differences observed in the tester and driver amplicons, a Southern analysis was performed on patient no. 14's tumor and matched normal genomic DNAs. Six μg of each DNA were cleaved with the *Msp*I enzyme and hybridized by probe *BR50*. The results showed that in the tumor DNA, the probe hybridized a lower band, ~1 kbp long, of much greater intensity than an upper band, which was ~2 kbp in length. In the normal DNA, just the opposite occurred (Fig. 2). Because the sizes of the upper and lower bands in the tumor and normal control DNAs were the same, the only reasonable explanation causing uneven hybridization intensities is DNA hypomethylation in the tumor cells. To examine whether the event also existed in other breast cancer patients, paired tumor and normal DNAs isolated from additional breast cancer biopsies were cleaved with *Msp*I and subjected to Southern analysis. The results showed that of 10 samples tested, 4 had similar hybridization patterns to those of patient no. 14 (Fig. 2). It is notable that in the normal DNAs of patient nos. 3, 4, and 5, more than one upper band was evident. We believe this can be attributed to partial DNA demethylation, instead of incomplete enzymatic digestion. This is because the Southern membrane was reblotted by a control probe, which was a background probe isolated along with probe *BR50*, and only a single hybridized band was displayed in each lane (Fig. 2).

***BR50* Also Detected Altered Methylation Patterns in Ovarian Cancer.** To examine whether the hypomethylation event also occurred in human ovarian cancer, paired DNA samples isolated from eight ovarian cancer patients were analyzed by probe *BR50*. The DNAs were cleaved with *Msp*I and hybridized with probe *BR50* in a Southern blot experiment. The results demonstrated that of eight patients tested, four displayed similar hybridization patterns to those observed in the breast cancer samples. In the tumor DNAs, the lower band was heavily hybridized by the probe, whereas in the normal control DNAs, a lower and upper band were hybridized. The completion of DNA digestion was confirmed by the same control probe used before, which only hybridized a single band in each lane (Fig. 3). Thus, *BR50* also detected altered DNA methylation in ovarian cancer.

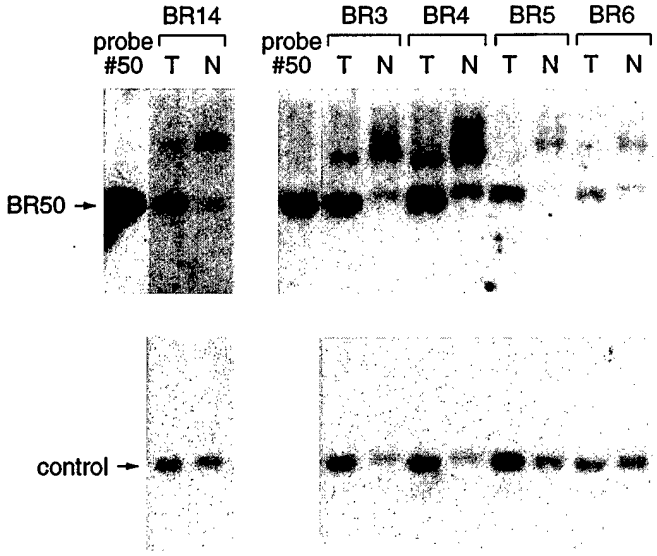


Fig. 2. The breast cancer genomic DNA Southern blot for probe *BR50*. *Left lane* of each blot, probe *BR50* hybridizes to itself as a positive control. *BR50* hybridizes with *Msp*I-digested 6 μg of original tumor (*T*) and matched normal (*N*) genomic DNAs isolated from patient no. 14 and other breast cancer patients. The number for each patient tested is listed above the *bracket*. Similar hypomethylation patterns in patient no. 14 and the other patients are observed. The control probe in the *lower section* served as an indicator of complete enzymatic digestion.

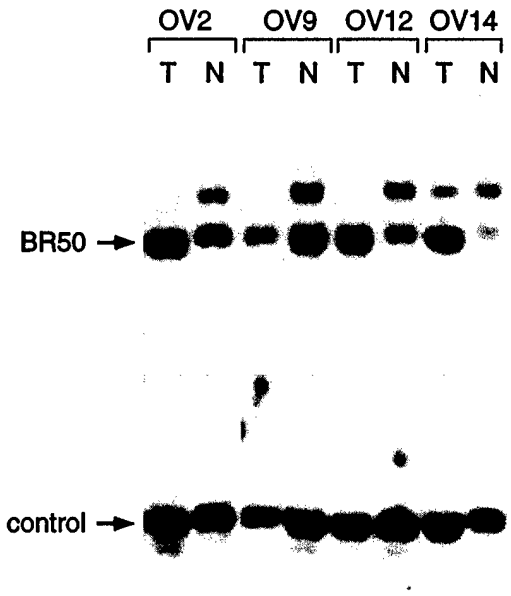


Fig. 3. The different methylation patterns displayed by probe *BR50* in ovarian cancer. The number for each patient is listed above the *bracket*. *Lanes T* and *N* include 6 μg of tumor DNA and normal DNA digested with *Msp*I, respectively. Tumor and matched normal DNAs were cleaved with *Msp*I and probed by *BR50* in a Southern blot experiment. Methyl-differential patterns are detected by *BR50*. The completion of the enzyme digestion was confirmed by the control probe, which only hybridized a single band in each lane.

Sequencing and Chromosome Assignment of Probe BR50. DNA sequencing found that *BR50* contains 1005 bps, with a CG content of 58% (Fig. 4*a*). The DNA fragment has a GGCCGG sequence on one end and CCGG sequence on the other end. Because *Msp*I is sensitive to GGC^{m}CGG and not sensitive to CC^{m}GG sequences, it is conceivable that the GGCCGG sequence was methylated in normal breast tissue DNA while being hypomethylated in tumor DNA. This prediction probably holds true because all of the candidate probes, including probe *BR254* and subsequent probes isolated from breast cancer biopsies, have one end terminating with a GGCCGG sequence and the other with a CCGG sequence (data not shown). A homologue search of the NIH GenBank discovered that *BR50* is not homologous to any existing sequences.[4] The chromosome assignment determined by monochromosomal human/rodent somatic cell hybrid mapping panel #2 established its location on chromosome 3 (data not shown). The fine chromosome position of *BR50* was analyzed by PCR amplification using GeneBridge 4 Radiation Hybrid Panel as templates. The top and bottom strand primers for *BR50* are 5′-ACCAGATGGAGGCAGTTGAC-3′ and 5′-AAGTGGGTGCTCTTTCTGTG-3′, respectively. The result obtained from radiation hybrid mapping suggests that *BR50* is placed 4.29 cR below the adjacent STS, AFMB362WB9, which is 179.84 cR from the top of the chromosome 3 linkage group. This result confirmed that *BR50* is proximally located on 3p12–14.

Search for the Coding Regions by Probe *BR50*. On the basis of our preliminary analysis, *BR50* was an interesting probe to be used for an adjacent gene search because it was differentially methylated in some breast and ovarian cancer samples tested, although the sample size was small. *BR50* is a genomic fragment, and consequently the chances that it encodes a polypeptide are slim because a majority of human genomic DNA sequences are noncoding sequences. As a result, we decided that the first step in searching for a gene should be to isolate a longer DNA piece from a human genomic DNA library,

[4] GenBank accession numbers: *BR50,* U78781; *TSP50,* AF100707.

a.

```
CCGGAGGCAGGCACAGGACTCGGGAGGGACGCTGCCAGCTCTCTGGGTGCTGAGTTCACAAGGCTG
CATTCATGATTTTCAATAGACCTGTGATGGTCTGTGCCCAGTGCTGGGGACACAGAAGAGTCAAAC
CTGGCTCCTGACCTGGACCTGGATCATCACGTGACAGGGAGGAGAGCGATCCAGGCTGATGAGGAA
AGCGCATGACATGGGGTCTTAGGAGCAGTGAGGGGCAGAGCCATGGCCAAAGGCCCCGCCATGGAA
GCTGAGGACTCTGGCACCAGATGGAGGCAGTTGACGGACCTCTGCCCTTGGGGTCCAACCCATGGG
CTTCTCATACATAGGGGTGAAAAAGGCCATTCTATTTATGCAGAATTTTCCCATGTGGCCAGGCAG
CAGAAGTCCAGAGGGGTAGGGGCCACTCAGGGTCACACAGAACAGCAGTTGCTGAAGACTGGGGAA
GTCCAGGCCTAGGCTCCACCTGCCCTTCCCCTGACATGGGGCCACCACTAGCCTTTTATGGGCAGG
CCTGGCTGCTGGTGGTTGGAATAACATCTGACTCCAGTGGGTGTCTGTCACCGTCTCCAGACAGGA
GACAGAGACAGAGGGTCAAAGTTCACTATGGCTCTTTGGGGCAATGAAATGCTGTGTTCTAGCCTC
TTGCCAGAAATCAGCCAAAGTCAAGGAAAGCCTGACTCCCACAGTTATCACAGAAAGAGCACCCAC
TTTCCAGCCCAGACAGCTGCACCCCAGCTGGGTCCTGGCAGCCCCAGCTTCAGCCTGGGCGGTATG
TTCCAGGCCCCTCGATCATCTGACCCTAATATCACCCCTTCACACCCCCTCCACTTTCTGCGGGAG
CCACCCCGAACCTTTGAATGGGGGAGATCCTGGAGGCTCTGCAATTTTCAGTGTAAACTGCCTGGA
GTTCCCCACTTCACCCTCATCTGGTTCACCTGTGGACTCCCAACAGAGCAGGCCCAGGAAACGCGG
GGCCTCTGAGGCCGG
```

b.

```
                                                    -59 gtcgtgggggcggc
    actgggagcgccttccggagagacgcagtcggctgccaccccggg

  1 atggggtcgctggtgccagaccgtcgcgcgcgggcagcgcccccgg
    M  G  R  W  C  Q  T  V  A  R  G  Q  R  P  R
 46 acgtctgccccctcccgcgccggtgccctgctgctgctgcttctg
    T  S  A  P  S  R  A  G  A  L  L  L  L  L  L
 91 ttgctgaggtctgcaggttgctggggcgcaggggaagccccgggg
    L  L  R  S  A  G  C  W  G  A  G  E  A  P  G
136 gcgctgtccactgctgatcccgccgaccagagcgtccagtgtgtc
    A  L  S  T  A  D  P  A  D  Q  S  V  Q  C  V
181 cccaaggccacctgtccttccagccggcctcgccttctctggcag
    P  K  A  T  C  P  S  S  R  P  R  L  L  W  Q
226 accccgaccacccagacactgccctcgaccaccatggagacccaa
    T  P  T  T  Q  T  L  P  S  T  T  M  E  T  Q
271 ttcccagtttctgaaggcaaagtcgacccataccgctcctgtggc
    F  P  V  S  E  G  K  V  D  P  Y  R  S  C  G
316 ttttcctacgagcaggaccccacccctcagggacccagaagccgtg
    F  S  Y  E  Q  D  P  T  L  R  D  P  E  A  V
361 gctcggcggtggccctggatggtcagcgtgcgggccaatggcaca
    A  R  R  W  P  W  M  V  S  V  R  A  N  G  T
406 cacatctgtgccggcaccatcattgcctcccagtgggtgctgact
    H  I  C  A  G  T  I  I  A  S  Q  W  V  L  T
451 gtggcccactgcctgatctggcgtgatgttatctactcagtgagg
    V  A  H  C  L  I  W  R  D  V  I  Y  S  V  R
496 gtggggagtccgtggattgaccagatgacgcagaccgcctccgat
    V  G  S  P  W  I  D  Q  M  T  Q  T  A  S  D
541 gtcccggtgctccaggtcatcatgcatagcaggtaccgggcccag
    V  P  V  L  Q  V  I  M  H  S  R  Y  R  A  Q
586 cggttctggtcctgggtgggccaggccaacgacatcggcctcctc
    R  F  W  S  W  V  G  Q  A  N  D  I  G  L  L
631 aagctcaagcaggaactcaagtacagcaattacgtgcggcccatc
    K  L  K  Q  E  L  K  Y  S  N  Y  V  R  P  I
676 tgcctgcctggcacggactatgtgttgaaggaccattcccgctgc
    C  L  P  G  T  D  Y  V  L  K  D  H  S  R  C
721 actgtgacgggctggggactttccaaggctgacggcatgtggcct
    T  V  T  G  W  G  L  S  K  A  D  G  M  W  P
766 cagttccggaccattcaggagaaggaagtcatcatcctgaacaac
    Q  F  R  T  I  Q  E  K  E  V  I  I  L  N  N
811 aaagagtgtgacaatttctaccacaacttcaccaaaatccccact
    K  E  C  D  N  F  Y  H  N  F  T  K  I  P  T
856 ctggttcagatcatcaagtcccagatgatgtgtgcggaggacacc
    L  V  Q  I  I  K  S  Q  M  M  C  A  E  D  T
901 cacagggagaagttctgctatgagctaactggagagcccttggtc
    H  R  E  K  F  C  Y  E  L  T  G  E  P  L  V
946 tgctccatggagggcacgtggtacctggtgggattggtgagctgg
    C  S  M  E  G  T  W  Y  L  V  G  L  V  S  W
991 ggtgcaggctgccagaagagcgaggccccacccatctacctacag
    G  A  G  C  Q  K  S  E  A  P  P  I  Y  L  Q
1036 gtctcctcctaccaacactggatctgggactgcctcaacgggcag
    V  S  S  Y  Q  H  W  I  W  D  C  L  N  G  Q
1081 gccctggccctgccagccccatccaggaccctgctcctggcactc
    A  L  A  L  P  A  P  S  R  T  L  L  L  A  L
1126 ccactgcccctcagcctccttgctgccctctgactctgtgtgccc
    P  L  P  L  S  L  L  A  A  L  *
1171 tccctcacttg
```

Fig. 4. *a*, the sequence of the genomic DNA probe *BR50*. *b*, nucleotide and predicted amino acid sequences of the human *TSP50* gene. The adenosine at the ATG (bold-type) initial codon is considered the number 1 nucleotide. The stop codon, TGA, is also in *boldface*. The sequences of exon *BR50-44* and *BR50-45* are in *italics*.

with the hope that it may contain exon(s). It is well known that DNA methylation sites can be near genes (18); therefore, we decided to screen a human placenta genomic phage library, EMBL3 SP6/T7 (Clontech, Inc.), where the insertion sizes were relatively small (9–28 kbp). For this purpose, 2×10^6 plaques derived from the library were screened by probe *BR50.* As a result, a 17-kbp length clone was isolated. To obtain the sequence information, the DNA clone was released from the phage DNA arms by restriction enzyme *Sst*I cleavage, which generated eight DNA fragments. All eight fragments were subcloned into pUC118 plasmids. The fragments smaller than 1 kbp were completely sequenced, whereas the fragments larger than 1 kbp were partially sequenced from both the plasmid and insert junctions. A homologue search of the NIH GenBank revealed two exons, *BR50-44* and *BR50-45,* which contained 112 and 132 nucleotides, respectively. Both exons encoded polypeptides that were ~50% identifiable to several mammalian proteases, such as serine proteases and tryptases. The *BR50-45* sequence was found 142 bp upstream of the *BR50* sequence.

Because we did not know the exact positions of the two exons in the 17-kbp fragment, it was possible that other exon(s) might lay between them. To gain a longer coding sequence, we designed four oligomers (A, B, C, and D), based on the sequence information of both exons, to perform PCR. Oligomer A (5′-CCTGGATGGTCAGCGTG-3′) and B (5′-CTGGGAGGCAATGATGGT-3′), which were on the complimentary strand, were based on the sequence information of *BR-44*; and C (5′-CTGGAGAGCCCTTGGTCT-3′) and D (5′-CAGTGTTGGTAGGAGGAG-3′), which were on the complimentary strand, were based on the sequence information of *BR-45.* A strategy using four different combinations of oligomer pairs was used to perform PCR by using the Human Universal cDNA Library Panel (Clontech, Inc.). A PCR product, which was about 700 bp in length, was generated from one oligomer combination (A/D, 5′-CCTGGATGGTCAGCGTG-3′/5′-CAGTGTTGGTAGGAGGAG-3′). This PCR product was directly sequenced. Combining the sequence information of the PCR product and the two exons, we obtained a cDNA fragment that contained 974 bps. The DNA homologue search of the NIH GenBank revealed again that it coded for a protease-like protein, and the overall identity was ~40%.

The Candidate Gene Is Highly and Specifically Expressed in Human Testes. To obtain a full-length cDNA, it is critical to use the right cDNA library where the gene of interest is expressed. Thus, two Human Multiple Tissue Northern blot panels, MTN and MTN II, containing 16 different tissue mRNAs (Clontech, Inc.), were used to test the expression of the candidate gene by using the 700-bp cDNA PCR product as a probe. The results showed that there were no visible transcripts of this gene in the eight mRNAs (heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas) included in the MTN panel (data not shown). In the MTN II panel, a 1.7-kbp band was heavily hybridized by the probe exclusively in the testes mRNA as compared with the control probe, which was the human *rab6* gene (Fig. 5). These results suggested that the gene that we were searching for is a tissue-specific gene. We have named the gene *TSP50.* At this moment, the biological function(s) of the gene in human testes remains unknown.

Isolation of the Full-length *TSP50* Gene from a Human Testes cDNA Library. *TSP50* is highly expressed in human testes; to search for the full-length *TSP50* gene, a human testes cDNA library (Human Testis cDNA Library; Clontech. Inc.) was screened by the 700-bp *TSP50* cDNA sequence. A cDNA clone containing the probe sequence was isolated. Sequence analysis suggests that this fragment encodes a protein with 385 animo acids (Fig. 4*b*). There is a stop codon located at the 117th bp upstream of the first initial translation site, and there is a 125-bp untranslated region before a polyadenosine

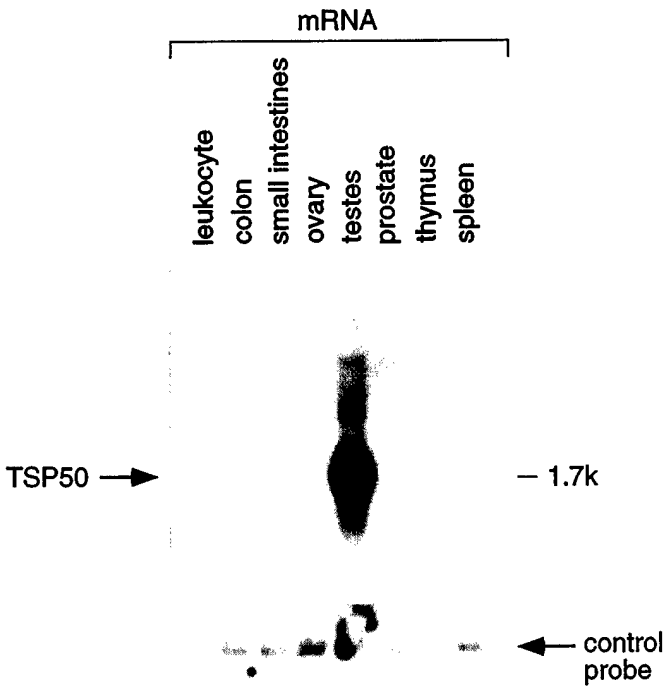


Fig. 5. Northern blot results analyzed by the 700-bp *TSP50* fragment. The Human Multiple Tissue Northern blot panel, MTN II, containing the mRNA of eight different tissues (Clontech Inc.), was tested to determine the expression levels of the gene. Compared with the control (the human *Rab6* gene, which was evenly expressed in all tissues), *TSP50* was highly expressed in the testes tissue but not in the other tissues.

sequence. These results imply that a full-length gene has been obtained. It is also notable that the *BR50* probe sequence is located at the 3′ end of the gene and is only 17 bp downstream from the polyadenosine adding signaling site. The exons *BR50-44* and *BR50-45* encode animo acids from 103 to 157 and 308 to 385, respectively (Fig. 4. b.). The 3′- untranslated region before the polyadenosine site is also included in the sequence of *BR50-45.*

DNA Methylation Status of the *TSP50* Gene in Human Testes and Other Normal Tissues. Our studies have proven that *TSP50* is a tissue-specific gene, and the methylation patterns in its 3′ region were altered in some breast and ovarian cancers. It is also known that many tissue specific genes are methylated, and this methylation may regulate their expression (22–24). To explore the possible relationship between *TSP50* gene expression and DNA methylation in different normal human tissues, Southern analysis was performed. The normal tissues tested included the testes, where *TSP50* was expressed, and bladder, blood, breast, colon, lung, kidney, placenta, and ovary samples, where *TSP50* was apparently not expressed. To perform the Southern analysis, *BR50* was used as a probe. DNAs isolated from the nine tissues were digested by *Msp*I and *Hpa*II, which is an isoschizomer of the *Msp*I enzyme and the most popular enzyme used to study DNA methylation patterns (25). *Hpa*II digestion showed that in the testes DNA, two bands, probably released from each allele by enzyme cleavage, were hybridized by the probe. However, in the DNAs of other tissues, the corresponding bands were either not hybridized or hybridized to a much smaller degree (Fig. 6*a*). For *Msp*I cleavage, both bands were released in different tissues to various extents (Fig. 6*b*). Both blots used a genomic fragment that did not detect differential DNA methylation as a control to determine complete enzymatic digestion (Fig. 6). These results demonstrated that the *TSP50* gene was differentially methylated in various human tissues. In general, DNA demethylation in the testes is correlated with high levels of gene expression. Conversely, DNA methylation is correlated with gene

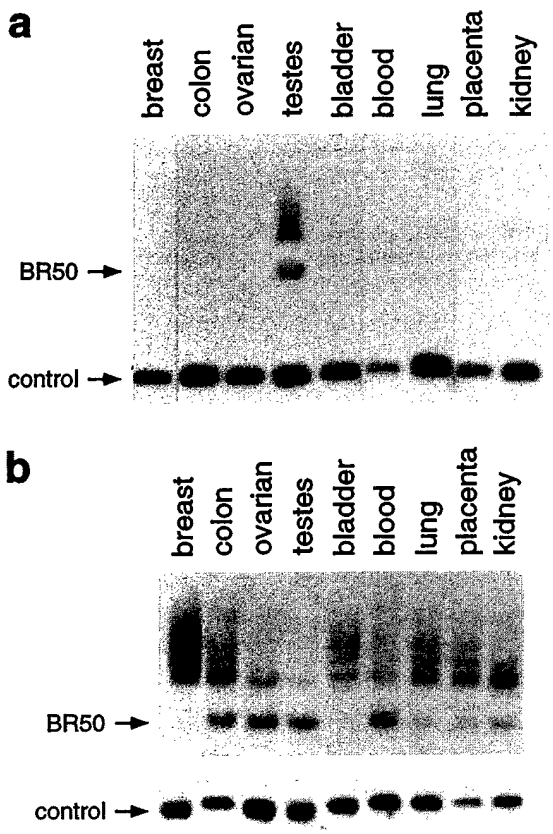


Fig. 6. DNA methylation status of the *TSP50* gene in nine normal human tissues examined by Southern blot. In *a* and *b*, the results obtained from *Hpa*II and *Msp*I digestion, respectively, are shown. Six μg of DNA isolated from each tissue were cleaved by the enzymes and subjected to Southern analysis by probe *BR50*. *a*, the results show that two bands, which are approximately 1 and 2 kbp in length, were released by *Hpa*II only in the testes tissue. *b*, a 2-kbp band was released by *Msp*I in most tissues. In *a* and *b*, the control probe hybridized a single band in the DNA of each tissue, which provides proof of complete enzymatic digestion.

silencing in the bladder, blood, breast, colon, lung, kidney, placenta, and ovary tissues.

TSP50 Was Differentially Expressed in Some Breast Cancer Tissues. Preliminary results demonstrated that *TSP50* was differentially methylated in 40% of the breast cancer tissues tested. This suggested that it could also be differentially expressed in cancer. To test this possibility, RT-PCR was carried out to determine *TSP50* expression levels in 18 paired breast cancer tissues. Our findings showed that *TSP50* PCR products were generated in five tumor tissues, whereas in their normal controls, they were not visible relative to the control gene, *β-actin* (Fig. 7). Products generated from the five patients were gel purified and sequenced. DNA sequence analysis confirmed that the PCR products synthesized were the *TSP50* gene (data not shown). Therefore, among the 18 paired samples tested, 28% of the tissues expressed the *TSP50* gene. At this moment, we cannot answer the question of whether activation of the *TSP50* gene in cancer is a consequence or a causal factor of neoplastic growth. To find the truth, it will be necessary to perform *in vitro* cellular transformation and *in vivo* tumor induction assays.

DISCUSSION

A modified RDA technique was used to study genetic alterations by using breast cancer biopsies as a working model system. As a result, two hypomethylated genomic DNA fragments were successfully isolated. The extensive study of one of the two fragments, *BR50*, is the subject of this report.

It has been reported that aberrant DNA methylations occur constantly in human tumors (9–12, 24, 26). DNA hypomethylations could activate oncogenes, whereas DNA hypermethylation could inactivate recessive oncogenes. Both events could result in neoplastic growth (27–35). The correlation between aberrant DNA methylations and malignancies suggests that differentially methylated fragments in tumors isolated by a modified RDA technique could be a valuable tool in the search for genes that might be related to cancer development. *BR50* was considered to hold such value because it not only detected DNA hypomethylations in the original breast cancer tissues from which it was isolated but also detected DNA hypomethylations in other breast and ovarian cancer samples.

Our first step in processing the gene search was to screen a human genomic phage library. A 17-kbp DNA fragment was isolated, and sequence analysis suggested that this fragment contained at least two exons that were homologous to mammalian proteases. The sequencing information of the exons led to the discovery of a 974-bp gene fragment from a human cDNA library panel by PCR amplification. To obtain a full-length gene, a Northern evaluation on 16 different types of human RNAs was performed. The results demonstrated that the target gene was specifically expressed in human testes tissue. This information secured the isolation of an intact gene, *TSP50*, by screening a human testes cDNA library. The sequence analysis revealed that the *TSP50* gene encodes a protein that shares ~40% identity with mammalian proteases, such as human tryptase or mouse serine protease. This would suggest that the product of the *TSP50* gene is a protease. However, at this point, we do not know the physiological function(s) of this protease. One may assume, though, that it could be a component in the human reproductive pathway due to it being solely expressed in the testes.

It is common knowledge that the expression of many tissue-specific genes is regulated by DNA methylations, which usually modify the promoter, or sometimes, 3′ regions (3, 4, 22–24, 36). Our preliminary results, although only obtained from analyzing the DNA methylation status of the 3′ flanking region of the gene, have proven that *TSP50* is one of those tissue-specific genes. It will be interesting to discover whether the promoter region of the gene is also methylated when the corresponding sequence information is available. The *Hpa*II and *Msp*I methylation-sensitive Southern analysis of the 3′ region of the *TSP50* gene demonstrated that, in *Hpa*II-digested DNAs, probe *BR50* hybridized two bands in the testes tissue but none in the other samples. The lower band, which was the same size as the probe, represented the unmethylated DNA pattern, whereas the upper band obviously con-

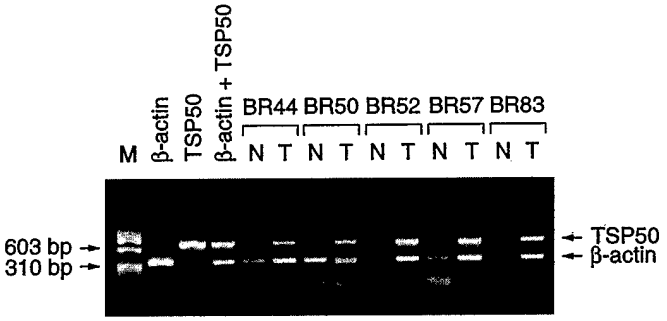


Fig. 7. The results of RT-PCR for *TSP50* differentially expressed in five breast cancer and normal control tissues. *Lane MW*, *Hae*III ø174 markers in bp. The *β-actin* and *TSP50* lanes served as positive controls; they were generated by using cDNA prepared from testes tissue RNA. *Lane β-actin+TSP50* contains simultaneously generated *TSP50* and *β-actin* from testes cDNA. The number for each patient tested is listed above the *bracket*. *T* and *N*, tumor and normal tissues, respectively.

tained the internal *Hpa*II recognition site(s), which remained methylated. In *Msp*I-digested DNAs, the upper band was dominant in most tissues, whereas in the testes, the lower band was dominant. These results suggest that the GGCCGG end of *BR50* was methylated in other tissues but not in the testes. The DNA methylation patterns observed in both blots are probably allelic orientated. It seems that DNA hypomethylation was accompanied by the expression of the gene in the testes, and conversely, DNA hypermethylation was accompanied by the silencing of the gene in other tissues. The correlation between DNA methylation and gene expression provided additional proof that DNA methylation could be an important mechanism in governing the expression of the genes in various differentiated human cells (12, 37, 38). In addition, the differential expression of the *TSP50* gene has been tested in 18 paired breast cancer biopsies. Our findings have shown that this gene was activated in five cancer samples. In the near future, more samples from different types of cancer will be examined, and the possibility that the *TSP50* gene product might be one of the factors that stimulate human cancer will be further explored.

Recently, by using the same technique, DNA fragments that represent DNA amplifications, deletions, and rearrangements were also obtained (data not shown). Hopefully, this technique will lead to the discovery of additional novel genes that may be related to cancer development. On the basis of our experience, the process of isolating the *TSP50* gene was made considerably easier by the modified technology, where the *Msp*I enzyme was used as the master enzyme. The ability of *Msp*I to recognize GC-rich sequences and its sensitivity to DNA methylation (17, 18) apparently accelerated our gene search. Furthermore, the double enzyme cleavage strategy provides another unique and efficient feature for this technique because ~40% of a human genome can theoretically be analyzed by a single master enzyme when it is combined with a different partner enzyme.

ACKNOWLEDGMENTS

We thank Craig Gawel for technical assistance and James Duffy for manuscript preparation.

REFERENCES

1. Shaw, D. J., Chaudhary, S., Rundle, S. A., Crow, S., Brook, J. D., Harper, P. S., and Harley, H. A study of DNA methylation in myotonic dystrophy. J. Med. Genet., *30:* 189–192, 1993.
2. Laird, P. W., and Jaenisch, R. DNA methylation and cancer. Hum. Mol. Genet., *3:* 1487–1485, 1994.
3. Riggs, D. A. 5-Methylcytosine, gene regulation, and cancer. *In:* J. P. Jost and H. P. Saluz (eds.), DNA Methylation: Molecular Biology and Biological Significance, pp. 1–30. Basel, Switzerland: Birkhauser Verlag, 1993.
4. Spruck, H. C., III, Rideout, M. W., III, and Jones, A. P. DNA methylation and cancer. *In:* J. P. Jost and H. P. Saluz (eds.), DNA Methylation: Molecular Biology and Biological Significance, pp. 487–509. Basel, Switzerland: Birkhauser Verlag, 1993.
5. Laird, P. W., Jackson-Grusby, L., Fazeli, A., Dickinson, S. L., Jung, W. E., Li, E., Weinberg, R. A., and Jaenisch, R. Suppression of intestinal neoplasia by DNA hypomethylation. Cell, *81:* 197–205, 1995.
6. Counts, J. L., and Goodman, J. I. Alteration in DNA methylation may play a variety of roles in carcinogenesis. Cell, *83:* 13–15, 1995.
7. Baylin, S. B., Herman, J. M., Graff, J. R., Vertino, P. M., and Issa, J-P. Alterations in DNA methylation. A fundamental aspect of neoplasia. Adv. Cancer Res., *72:* 141–196, 1998.
8. Bird, P. A. CpG islands as gene markers in the vertebrate nucleus. Trends Genet., *3:* 342–346, 1987.
9. Larsen, F., Gunderson, G., Lopez, R., and Prydz, H. CpG island as gene markers in the human genome. Genomics, *13:* 1095–1107, 1992.
10. Adams, R. L., and Buron, R. H. DNA methylation in eukaryotes. CRC Crit Rev. Biochem., *13:* 349–384, 1982.
11. Antequera, F., and Bird, P. A. CpG islands. *In:* J. P. Jost and H. P. Saluz (eds.), DNA Methylation: Molecular Biology and Biological Significance, pp. 169–185. Basel, Switzerland: Birkhauser Verlag, 1993.
12. Bird, P. A. CpG-rich islands and the function of DNA methylation. Nature (Lond.), *321:* 209–213, 1986.
13. Gardiner, K. Base composition and gene distribution: critical patterns in mammalian genome organization. Trends Genet., *12:* 519–524, 1996.
14. Lisitsyn, N., Lisitsyn, N., and Wigler, M. Cloning the differences between two complex genomes. Science (Washington DC), *259:* 946–951, 1993.
15. Cedar, H., Solage, A., Glaser, G., and Razin, A. Direct detection of methylated cytosine in DNA by use of the restriction enzyme. Msp I. Nucleic Acids Res., *6:* 2125–2133, 1979.
16. Busslinger, M., deBoer, E., Wright, S., Grosveld, F. G., and Flavell, R. A. The sequence GGC^{m}CGG is resistant to *Msp*I cleavage. Nucleic Acids Res., *11:* 3559–3569, 1983.
17. Antequera, F., Boyes, J., and Bird, A. High levels of *de novo* methylation and altered chromatin structure at CpG islands in cell lines. Cell, *62:* 503–514, 1990.
18. Weng, A., Engler, P., and Storb, U. The bulk chromatin structure of a murine transgene does not vary with its transcriptional or DNA methylation status. Mol. Cell. Biol., *15:* 572–579, 1995.
19. Bishop, D. T., Williamson, A. J., and Skolnick, H. M. A model for restriction length distributions. Am. J. Hum. Genet., *35:* 795–815, 1983.
20. Sambrook, J., Fritsch, E. F., and Maniatis, T. *In:* N. Ford, C. Nolan, and M. Ferguson (eds.), Molecular Cloning Manual, pp. 9.16–9.19. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1989.
21. Walter, M. A., Spillett, D. J., Thomas, P., Weissenbach, J., and Goodfellow, P. N. A method for constructing radiation hybrid maps of whole genomes. Nat. Genet., *7:* 22–28, 1994.
22. Bird, A. Gene number, noise reduction and biological complexity. Trends Genet., *11:* 94–100, 1995.
23. Kass, S. U., Pruss, D., and Wolffe, A. P. How does DNA methylation repress transcription? Trends Genet., *13:* 444–449, 1997.
24. Siegfried, Z., and Cedar, H. DNA methylation: a molecular lock. Curr. Biol., *7:* R305–R307, 1997.
25. Nelson, M., and McClelland, M. Site-specific methylation: effect on DNA methylation methyltransferases and restriction endonucleases. Nucleic Acids Res., *19:* 2045–2071, 1991.
26. Yeivin, A., and Razin, A. Gene methylation patterns and expression. *In:* J. P. Jost and H. P. Saluz (eds.), DNA Methylation: Molecular Biology and Biological Significance, pp. 523–568. Basel, Switzerland: Birkhauser Verlag, 1993.
27. Cheah, M. S., Wallace, C. D., and Hoffman, R. M. Hypomethylation of DNA in human cancer cells: a site-specific change in the c-*myc* oncogene. J. Natl. Cancer Inst., *73:* 1057–1065, 1984.
28. Hanada, M., Delia, D., Aiello, A., Stadtmauer, E., and Reed, J. C. *bcl-2* gene hypomethylation and high-level expression in B-cell chronic lymphocytic leukemia. Blood, *82:* 1820–1828, 1993.
29. Ray, J. S., Harbison, M. L., McClain, R. M., and Goodman, J. I. Alterations in the methylation status and expression of the *raf* oncogene in phenobarbital-induced and spontaneous B6C3F1 mouse liver tumor. Mol. Carcinog., *9:* 155–166, 1994.
30. Vorce, R. L., and Goodman, J. I. Altered methylation of *ras* oncogenes in benzidine-induced B6C3F1 mouse liver tumors. Toxicol. Appl. Pharmacol., *100:* 398–410, 1989.
31. Herman, J. G., Latif, F., Weng, Y., Lerman, M. I., Zbar, B., Liu, S., Samid, D., Duan, D-S. R., Gnarra, J. R., Linehan, M., and Baylin, S. B. Silencing of the *VHL* tumor-suppressor gene by DNA methylation in renal carcinoma. Proc. Natl. Acad. Sci. USA, *91:* 9700–9704, 1994.
32. Merlo, A., Herman, J. G., Mao, L., Lee, D. J., Gabrielson, E., Burger, P. C., Baylin, S. B., and Sidransky, D. 5′ CpG island methylation is associated with transcriptional silencing of the tumor suppressor *p16/CDKN2/MTS1* in human cancers. Nat. Med., *1:* 686–692, 1995.
33. Herman, J. G., Jen, J., Merlo, A., and Baylin, S. B. Hypermethylation-associated inactivation indicates a tumor suppressor role for p15^{INK4B1}. Cancer Res., *56:* 722–727, 1996.
34. Graff, J. R., Herman, J. G., Lapidus, R. G., Chopra, H., Xu, R., Jarrard, F., Isaacs, W. B., Pitha, P. M., Davidson, N. E., and Baylin, S. B. E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. Cancer Res., *55:* 5195–5199, 1995.
35. Baylin, S. B., Herman, J. G., Graff, J. R., Vertino, P. M., and Issa, J-P. Alterations in DNA methylation: a fundamental aspect of neoplasia. Adv. Cancer Res., *72:* 141–196, 1998.
36. Graff, J. R., Herman, J. G., Myohanen, S., Baylin, S. B., and Vertino, P. M. Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in *de novo* methylation. J. Biol. Chem., *272:* 22322–22329, 1997.
37. Razin, A., and Kafri, T. DNA methylation from embryo to adult. Prog. Nucleic Acid Res. Mol. Biol., *48:* 53–58, 1994.
38. Martienssen, R. A., and Richards, E. J. DNA methylation in eukaryotes. Curr. Opin. Genet. Dev., *5:* 234–242, 1995.